

## Construction and characterization of an infectious molecular clone of HIV-1 subtype A of Indian origin

Milka A. Rodriguez<sup>a</sup>, Yue Chen<sup>a</sup>, Jodi K. Craig<sup>b</sup>, Ramdas Chatterjee<sup>c</sup>, Deena Ratner<sup>a</sup>, Masashi Tatsumi<sup>d</sup>, Pratima Roy<sup>e</sup>, Dhruva Neogi<sup>e</sup>, Phalguni Gupta<sup>a,\*</sup>

<sup>a</sup> Department of Infectious Diseases and Microbiology, Graduate School of Public Health, University of Pittsburgh, 426 Parran Hall, 130 DeSoto Street, PA 15261, USA

<sup>b</sup> Department of Molecular Genetics and Biochemistry, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15261, USA

<sup>c</sup> Department of Virology, Chittaranjan Cancer Research Institute, Calcutta, India

<sup>d</sup> Department of Veterinary Science, National Institute of Infectious Diseases, Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

<sup>e</sup> Department of Virology, School of Tropical Medicine, Calcutta, India

Received 2 June 2005; returned to author for revision 9 August 2005; accepted 21 September 2005

Available online 9 November 2005

### Abstract

India has the second highest number of HIV-1 infected people next to South Africa. The predominant proportion of HIV-1 circulating in India is of subtype C origin, with a small fraction made up of subtypes A and B. In this report, we describe the construction and characterization of the first full-length infectious molecular clone p1579A-1 HIV-1, from an HIV-1 subtype A infected person from India, using long PCR and successive ligation of the amplimers. Phylogenetic analysis of the sequence of the entire proviral DNA and LTR confirmed p1579A-1 to be an HIV-1 subtype A. Analysis of the *env* gene of p1579A-1 showed a conserved GPGQ motif and the absence of basic amino acids at positions 11 and 25 suggesting CCR5 coreceptor usage. Analysis of *env* N-linked glycosylation sites revealed fewer sites in the V1 region of envelope compared to other subtype A. Transcription factor binding site analysis of the LTR sequences identified conserved as well as unique transcription factor binding sites (TFBS) in p1579A-1. This infectious clone of HIV-1 can be useful to study the molecular mechanism of dominance of subtype C in India.

© 2005 Elsevier Inc. All rights reserved.

**Keywords:** HIV-1; Molecular clone; Subtype A; India

HIV-1 was first reported in Tamil Nadu, India, in 1986 among female commercial sex workers (John et al., 1987; Simoes et al., 1987). Since then, HIV-1 has spread rapidly throughout the country and has been reported in nearly every major city. Today, India has the second largest number of HIV-1 infected individuals in the world next to South Africa (UNAIDS and W.H.O., 2004). Genetic analyses of HIV-1 sequences circulating in different parts of India have shown that the predominant proportion of HIV-1 circulating in India is of subtype C origin with a small fraction of subtypes A and B (Cassol et al., 1996; Delwart et al., 1998; Maitra et al., 1999; Mandal et al., 2000, 2002; Shankarappa et al., 2001; Tripathy et al., 1996). Furthermore, phenotypic and genotypic analyses of *env* sequences of HIV-1 in India indicate that they are CCR5

tropic with non-syncytium inducing phenotype even when they are isolated from late stage of infection (Cecilia et al., 2000; Shankarappa et al., 2001). We have recently analyzed subtype C sequences from India and then compared this set of sequences to subtype C sequences sampled from Botswana, Burundi, South Africa, Tanzania, and Zimbabwe. Overall, HIV-1 type C sequences from different parts of India were more closely related to each other (10%) than to subtype C sequences from Botswana, Burundi, South Africa, Tanzania, and Zimbabwe (15–21%) (Shankarappa et al., 2001). These results indicate that subtype C sequences in India are distinct from subtype C sequences sampled from other countries.

Over the course of the HIV-1 epidemic in India, and regardless of geographical area, subtype C has consistently accounted for 90–95% of infections with subtypes A, B, and others accounting for the remaining 5–10% (Cassol et al., 1996; Delwart et al., 1993; Jameel et al., 1995; Tripathy et al.,

\* Corresponding author. Fax: +1 412 624 4953.

E-mail address: [pgupta1@pitt.edu](mailto:pgupta1@pitt.edu) (P. Gupta).

1996; Tsuchie et al., 1995). This together with the fact HIV-1 isolated from different parts of India at different times are closely related (Shankarappa et al., 2001) suggests that the preponderance of subtype C viruses over other subtypes is most probably not due to new and continual introductions of HIV-1 subtype C into the country or to recent immigration or representative of a cluster of isolated individuals. It is possible that the disproportionate distribution of HIV-1 subtypes in India may be due to one or a combination of factors including founder effect, replication fitness, and transmission efficiency. In order to elucidate the molecular dynamics of HIV-1 subtype distribution in India, it is important to have an infectious molecular clone, such as subtype A, from India. Reports of subtype A viruses in India date back to 1991 (Cassol et al., 1996) and since then they have been periodically reported in various studies (Ljungberg et al., 2002; Sahni et al., 2002; Shankarappa et al., 2001). To date, there are no sequence data of the complete genome of an HIV-1 subtype A nor does a replication competent infectious molecular clone of subtype A exist. Here, we report the construction of a full-length infectious molecular clone of subtype A HIV-1 from India and delineation of its complete genomic sequence.

## Results and discussion

From the molecular epidemiological data, it is expected that very few subjects will be infected with subtype A HIV-1. Therefore, we have performed virus isolation from a number of HIV-1 infected subjects and determined their subtype specificity. The infected subjects for this study were recruited from Calcutta, the major eastern city of India. A summary of patient characteristics is shown in Table 1. HIV-1 was isolated from

these patients by coculturing their PBMC with PHA-stimulated CD8 depleted normal donor PBMC as described previously (Balachandran et al., 1991; Balachandran et al., 1988). Production of virus was monitored by HIV-1 p24 production in the culture supernatant. Following 30 to 40 days of cultivation, chromosomal DNA was extracted from the infected cell pellets. The DNA was then used to determine the subtypes of HIV-1 isolates using a heteroduplex mobility assay (HMA) kit obtained through the NIH AIDS Repository (Bachmann et al., 1994; Delwart et al., 1993, 1995). In this assay, a nested PCR reaction was employed which in the first round amplified a ~1.25-kb-long amplicon encompassing the V1–V5 region of the HIV-1 envelope gene from proviral DNA. The 1.25-kb DNA was then used as a template in the second round of PCR to amplify a ~0.7-kb C2–V5 envelope gene fragment. Similar nested PCR primer pairs were also used for amplification of the C2–V5 region from reference plasmids carrying HIV-1 subtype A, B, and C env sequences. Nested PCR products from each of the reference plasmids and the unknown were then combined, denatured and re-annealed. Subtypes were assigned based on the mobility of heteroduplexes on a non-denaturing polyacrylamide gel. Among the 19 HIV-1 positive Indian isolates tested, one, 1579, was found to be a subtype A (Table 1). Subtypes were further verified by sequencing the 0.7-kb PCR products followed by its blast analysis at the Los Alamos HIV-1 Database (data not shown). Proviral DNA from this isolate (1579A) was then used to construct the subtype A infectious molecular clone.

To construct the full-length molecular clone, long PCR was performed using a high fidelity polymerase and chromosomal DNA from HIV-1 1579A infected PBMC. Three overlapping subgenomic fragments of 2.5 kb: 5′LTR-pol, 6.2 kb: gag-env, and 2.7 kb: env-3′LTR were generated. These three fragments were cloned into pCR-Blunt II-Topo cloning vectors (Invitrogen) resulting in plasmids A, B, and C, respectively (Fig. 1). Plasmids A and B were digested with unique restriction enzymes AhdI and FseI, and the resulting fragments were ligated to generate plasmid D containing a 7.6-kb fragment of 1579A genome. Plasmids C and D were then digested with unique restriction enzymes DraIII and SnaBI, and the resulting fragments were ligated to generate the full-length infectious molecular clone, p1579A. Our cloning strategy exploited the pUC origin of replication and kanamycin resistance gene within the cloning vector such that the fragments being ligated together contained either one or the other of these features. This technique significantly lowered the background of our ligation products and facilitated cloning of these large fragments.

To validate the presence of the entire HIV-1 genome in p1579A clones, PCR was performed on three separate clones using gag, pol, and env gene specific primers. Each of the three clones produced amplicons of correct size, 0.9 kb, 1 kb, and 0.7 kb, corresponding to HIV-1 gag, pol and env genes, respectively (Fig. 2). The intactness of these three clones gained further support by the analysis of the digestion pattern with NotI restriction enzyme which cleaves outside of the HIV-1 subtype A genome on either side. An approximately 10-kb band corresponding to full-length HIV-1 DNA and a 3.8-kb band corresponding to the plasmid vector were observed for

Table 1  
Characteristics of HIV-1 infected subjects from India

Patient ID#	Age	Sex	Clinical Symptoms	CD4 Count	Subtype
2145	33	F	Asymptomatic	678/ul	ND
1577	18	F	Asymptomatic	814	C
2161	53	M	Asymptomatic	427	ND
1590	16	F	Weakness, anorexia	736	ND
2024	25	F	Asymptomatic	350	C
1310	25	F	Abdomen pain, diarrhea	343	C
1581	21	F	Fever, cough, weakness	377	C
1580	18	F	Asymptomatic	810	C
1540	20	F	Total body pain, menstrual problems	640	ND
1579	16	F	Asymptomatic	477	A
1443	34	F	Weakness	702	C
2167	25	F	Abdomen pain, leucorrhoea	255	C
1585	21	F	Weakness	331	C
2163	25	F	Weakness, anorexia	275	C
2177	18	F	Cold and cough, weakness	193	C
1583	18	F	Asymptomatic	945	C
1521	30	F	Weakness, blood pressure 31/87	221	ND
2114	ND	ND	Asymptomatic	ND	C
1451	26	F	Headache, cough, weight loss, weakness	291	ND

ND—not determined.

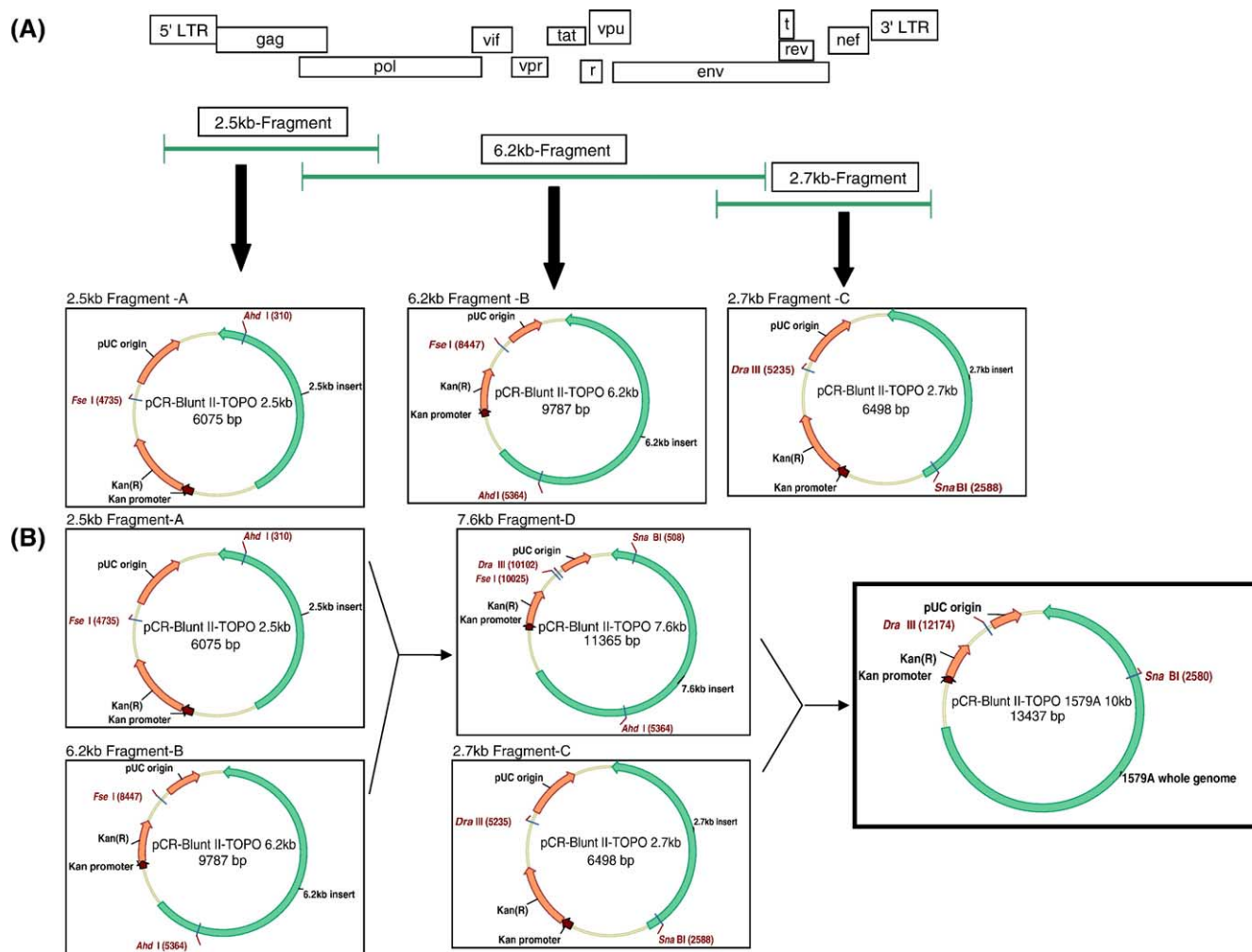


Fig. 1. Construction of an HIV-1 subtype A infectious molecular clone. (A) PCR and cloning of three overlapping subgenomic fragments of 1579A genome. (B) Cloning strategy. Each of the plasmids was cut with unique restriction enzymes and the products ligated together to recover a complete full-length clone of 1579A.

each clone that was digested (data not shown). One of these three clones, p1579A-1, was used in subsequent biologic and genetic studies.

The infectivity of p1579A-1 was examined by transfection of human endothelial kidney (HEK) 293T cells. Forty eight

hours after transfection, filtered culture supernatant was used to infect CD8 depleted PBMC (Kalia et al., 2003). Virus growth was monitored by measuring HIV-1 p24 antigen in culture supernatant. An infectious molecular clone of HIV-1 subtype C of Indian origin, pIndieC1 (Mochizuki et al., 1999) was used as a positive control for transfection. Results shown in Fig. 3A indicate that p1579A-1 is able to produce replication competent HIV-1.

To characterize the growth properties of HIV-1 p1579A-1, CD8 depleted PBMC were infected with p1579A-1 virus (equivalent to 21ng of p24) collected from transfected culture supernatant and an equivalent p24 value of the parental HIV-1 subtype A isolate. Kinetics of viral replication was determined by measuring HIV-1 p24 antigen in culture supernatant. Fig. 3B shows that p1579A-1 cloned virus has similar growth properties as the parental isolate. However, the parental isolate produced slightly higher p24 values than p1579A-1.

To determine the coreceptor usage of p1579A-1 cloned virus, U87.CD4 cells expressing either the CXCR4 or CCR5 coreceptors were infected with virus containing supernatant equivalent to 30ng of HIV-1 p24 antigen. Virus growth was

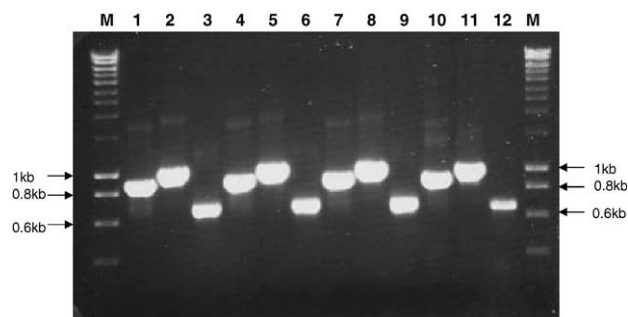


Fig. 2. Validation of whole genome clone using HIV-1 PCR primers for gag, pol and env genes. M—marker. Lanes 1, 2, 3; gag, pol, env (respectively) PCR products for clone #1. Lanes 4, 5, 6; gag, pol, env (respectively) PCR products for clone #2. Lanes 7, 8, 9; gag, pol, env (respectively) PCR products for clone #3. Lanes 10, 11, 12; gag, pol, env (respectively) PCR products for positive control of chromosomal DNA from 1579-infected PBMC.

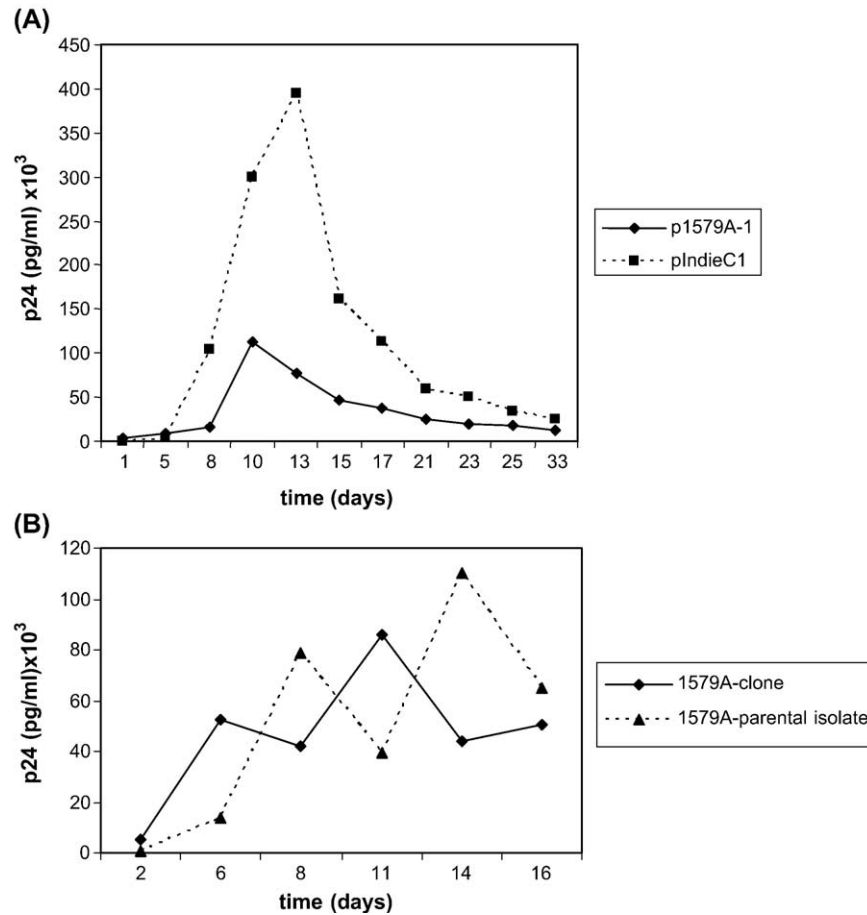


Fig. 3. (A) Transfection-infection of p1579A-1 DNA. HEK293T cells were first transfected with p1579A-1 or the control pIndieC1 DNA. Two days after transfection, culture supernatant was then used to infect CD8 depleted PBMC. Production of HIV-1 p24 antigen on each day sampled is shown for p1579A-1 and the control pIndieC1. (B) Replication kinetics of p1579A-1. CD8 depleted PBMC were infected with 21ng of HIV-1 p24 equivalent virus supernatant of p1579A-1 or the parental isolate, 1579A. Virus growth was monitored by measuring HIV-1 p24 antigen in the culture supernatant every 2 days.

monitored by measuring p24 antigen in culture supernatant. p1579A-1 HIV-1 replicated more efficiently in U87.CD4 cells expressing CCR5 than CXCR4, indicating that p1579A-1 virus is an R5 tropic virus (Table 2). Consistent with CCR5 coreceptor usage, p1579A-1 HIV-1 was found to be non-syncytia inducing (NSI), due to its inability to form syncytia in MT2 cells (data not shown).

Table 2  
Coreceptor usage of p1579A-1 cloned virus

Time after infection (days)	HIV-1		BAL <sup>b</sup>		IIB <sup>c</sup>	
	p1579A-1 <sup>a</sup>		CCR5	CXCR4 <sup>d</sup>	CCR5 <sup>e</sup>	CXCR4
2	183	276	2,276	601	338	>200,000
4	705	210	99,275	650	425	>200,000
6	3480	220	ND	ND	ND	ND
8	12,770	280	ND	ND	ND	ND
10	33,625	300	ND	ND	ND	ND

<sup>a,b,c</sup> HIV-1 p24 antigen production (pg/ml) in U87.CD4.CCR5 or U87.CD4.CXCR4 expressing cells.

<sup>d,e,f</sup> An increase in p24 value of at least 3-fold over the day 2 value is considered positive HIV replication.

ND—not determined.

The complete sequence of the proviral DNA of p1579A-1 was determined by primer walking along the entire genome. The sequences were verified using the HIV-1 Sequence Locator tool at the Los Alamos HIV-1 sequence database and then manually joined. The full-length genome of p1579A-1 is 9699-bp long. All reading frames of this clone were found to be open. The entire genome was subjected to HIV-1 recombination analysis at the Los Alamos Sequence Database and was not found to be an HIV-1 recombinant. The complete genome of p1579A-1 and the long terminal repeat (LTR) were aligned with HIV-1 group M subtype reference sequences obtained from the Los Alamos HIV-1 database. Phylogenetic analysis of both the complete genome and the LTR showed that 1579A clusters within the HIV-1 subtype A lineage (Fig. 4).

The envelope gene is important for viral attachment, infectivity, and coreceptor usage. We analyzed the envelope gene by aligning the predicted amino acid sequence of p1579A-1 with those of the consensus sequence of reference subtypes A1 and A2 and pIndieC1. The characteristic GPGQ motif at the crown of the V3 loop was conserved in all sequences (Fig. 5). Consistent with earlier findings of CCR5 coreceptor usage and an NSI phenotype, there was an absence of basic amino acids at positions 11 and 25 in p1579A-1 (Fig. 5). This was also true for

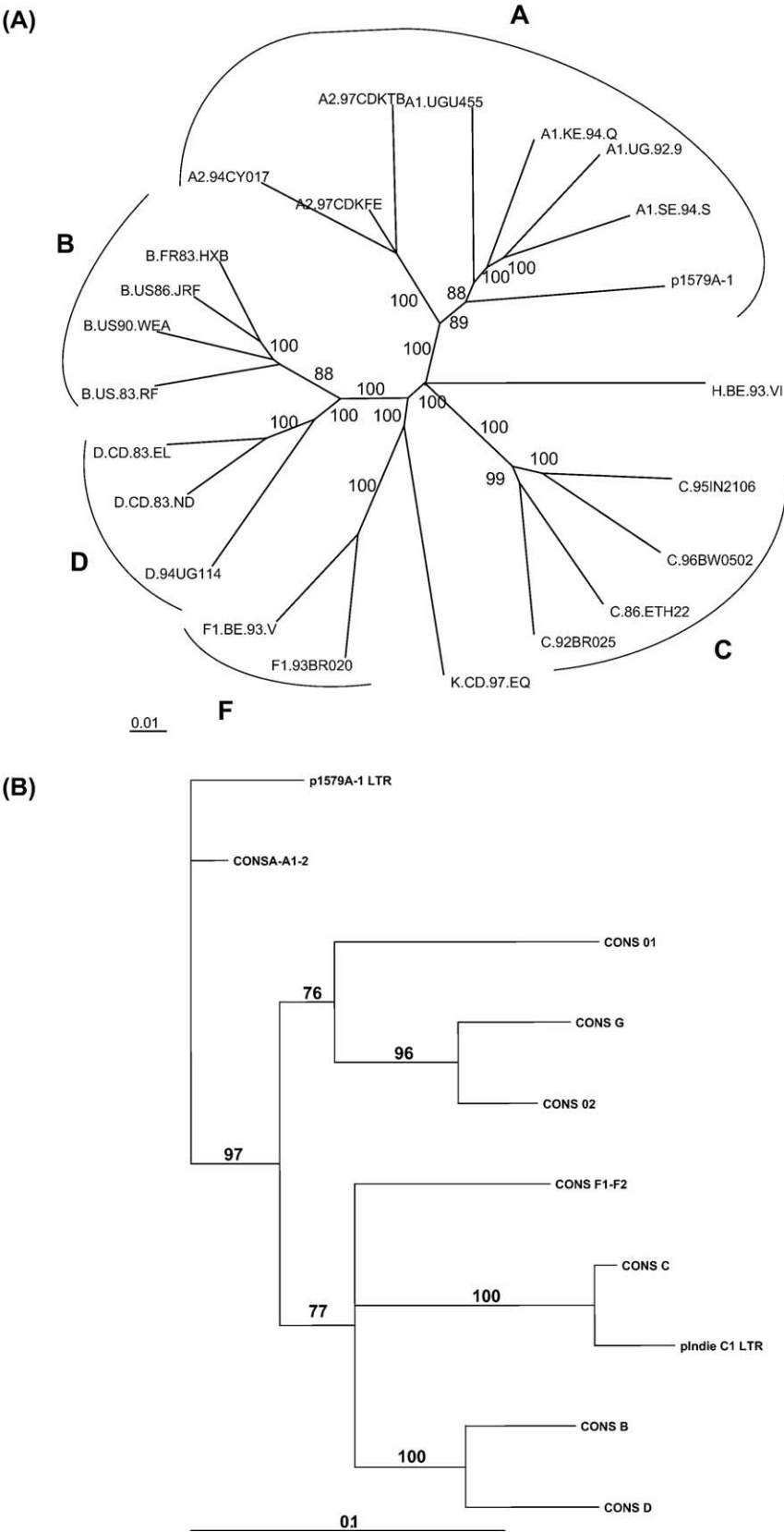


Fig. 4. Phylogenetic analysis of p1579A-1. (A) Neighbor-joining Jukes Cantor unrooted tree showing the phylogenetic relationship of the complete genomic sequence of p1579A-1 to HIV-1 group M reference subtypes. (B) Maximum likelihood unrooted tree showing the phylogenetic relationship of p1579A-1 long terminal repeat (LTR) sequence to the consensus (CONS) sequence of HIV-1 group M LTR. All HIV-1 group M reference sequences were obtained from the Los Alamos HIV database.



1579A	mrvmetqrqsy	qhlwrgwtmi	lgmliisaa	enlwvtvyyg	vpvkwdaett	lfcasdakay	etekhvwat	hacvptdlnp	qeihlkNvte	dfnmwnnmv
CONSENSUS_A1	mrvmgigqnc	qhlwrgwtmi	lgmliisaa	enlwvtvyyg	vpvkwdaett	lfcasdakay	etemhvwat	hacvptdnp	qeihleNvte	efnmwnnmv
CONSENSUS_A2	mrvmgigqny	qhlwrgwtmi	lgmliisaa	enlwvtvyyg	vpvkwdaett	lfcasdakay	dtevhvwat	hacvptdnp	qeivleNvte	dfnmwnnmv
Indie_C1	mrvrgtllrny	qgwviwvlg	fwlmimcngg	gnlwvtvyyg	vpvkwdaett	lfcasdakay	erevhvwat	hacvptdnp	qeivlgNvte	nfnmwnnmv
V1										
1579A	eqmhtdiisl	wdgslkpcve	ltplcvtlnc	snvNnt---	-----gev	kNctfNmtte	lrdkrqkvss	lfyrldivpi	kqgggNraNs	syeyrllnc
CONSENSUS_A1	eqmhtdiisl	wdgslkpcvk	ltplcvtlnc	snvNntNntt	-----ntheei	kNcsfNmtte	lrdkkqkvys	lfyrldivpi	ne---NnaNs	---syrlinc
CONSENSUS_A2	eqmhtdiisl	wdgslkpcvk	ltplcvtlnc	snvNntNntt	-----ntheei	kNcsfNmtte	lrdkkqkvys	lfyrldivpi	ne---NnaNs	---syrlinc
Indie_C1	dqmhedvisl	wdgslkpcvk	ltplcvtlnc	snvNntNntt	-----ntheei	kNcsfNmtte	lrdkkqkvys	lfyrldivpi	ne---NnaNs	---syrlinc
V2										
1579A	Ntsaitqacp	kvtfepipih	yctpagfail	kckdeefNgt	gpcrNvstvg	cthgikpvvs	tqlllNgsla	knkviiraeN	itnnvktiv	qlaepviinc
CONSENSUS_A1	Ntsaitqacp	kvtfepipih	yctpagfail	kckdeefNgt	gpcrNvstvg	cthgikpvvs	tqlllNgsla	knkviiraeN	itnnvktiv	qlaepviinc
CONSENSUS_A2	Ntsaitqacp	kvtfepipih	yctpagfail	kckdeefNgt	gpcrNvstvg	cthgikpvvs	tqlllNgsla	knkviiraeN	itnnvktiv	qlaepviinc
Indie_C1	Ntsaitqacp	kvtfepipih	yctpagfail	kckdeefNgt	gpcrNvstvg	cthgikpvvs	tqlllNgsla	knkviiraeN	itnnvktiv	qlaepviinc
V3										
1579A	trpnNntres	irgpggdfy	ts-niigdir	qahcNwksa	wNktlhavge	glekyfvNkt	ikfdrptggd	leitthsfnc	ggeffycNts	elfNstwnss
CONSENSUS_A1	trpnNntres	irgpggdfy	ts-niigdir	qahcNwksa	wNktlhavge	glekyfvNkt	ikfdrptggd	leitthsfnc	ggeffycNts	elfNstwnss
CONSENSUS_A2	trpnNntres	irgpggdfy	ts-niigdir	qahcNwksa	wNktlhavge	glekyfvNkt	ikfdrptggd	leitthsfnc	ggeffycNts	elfNstwnss
Indie_C1	trpnNntres	irgpggdfy	ts-niigdir	qahcNwksa	wNktlhavge	glekyfvNkt	ikfdrptggd	leitthsfnc	ggeffycNts	elfNstwnss
V4										
1579A	tqgt---Ntel	Ndtitlpcr	kqvinmwqr	gqamyappik	gliicvsnit	gilltrdgg	Ns---tNgt	tfrpgggdmr	dnwrselyky	kvvqieplgv
CONSENSUS_A1	-----Ngtm	Ndtitlpcr	kqvinmwqr	gqamyappik	gliicvsnit	gilltrdgg	Ns---tNgt	tfrpgggdmr	dnwrselyky	kvvqieplgv
CONSENSUS_A2	-----Ngtm	Ndtitlpcr	kqvinmwqr	gqamyappik	gliicvsnit	gilltrdgg	Ns---tNgt	tfrpgggdmr	dnwrselyky	kvvqieplgv
Indie_C1	-----Ngtm	Ndtitlpcr	kqvinmwqr	gqamyappik	gliicvsnit	gilltrdgg	Ns---tNgt	tfrpgggdmr	dnwrselyky	kvvqieplgv
V5										
1579A	aptkarrtrv	grekraigl	avflgflgaa	gstmgaaait	ltvqarqlls	givqgqsnll	raieaqqgml	rltwgikql	qarvlaery	lrdqgllgwi
CONSENSUS_A1	aptkarrtrv	grekraigl	avflgflgaa	gstmgaaait	ltvqarqlls	givqgqsnll	raieaqqgml	rltwgikql	qarvlaery	lrdqgllgwi
CONSENSUS_A2	aptkarrtrv	grekraigl	avflgflgaa	gstmgaaait	ltvqarqlls	givqgqsnll	raieaqqgml	rltwgikql	qarvlaery	lrdqgllgwi
Indie_C1	aptkarrtrv	grekraigl	avflgflgaa	gstmgaaait	ltvqarqlls	givqgqsnll	raieaqqgml	rltwgikql	qarvlaery	lrdqgllgwi
V6										
1579A	gcsqklictt	nvpwNsswn	ktqeeliwNtm	twlqwdkeis	Nytliliynli	eesqngqekn	ekdlldldkw	atlwswfdit	kwlywikifi	mivggliglr
CONSENSUS_A1	gcsqklictt	nvpwNsswn	ktqeeliwNtm	twlqwdkeis	Nytliliynli	eesqngqekn	ekdlldldkw	atlwswfdit	kwlywikifi	mivggliglr
CONSENSUS_A2	gcsqklictt	nvpwNsswn	ktqeeliwNtm	twlqwdkeis	Nytliliynli	eesqngqekn	ekdlldldkw	atlwswfdit	kwlywikifi	mivggliglr
Indie_C1	gcsqklictt	nvpwNsswn	ktqeeliwNtm	twlqwdkeis	Nytliliynli	eesqngqekn	ekdlldldkw	atlwswfdit	kwlywikifi	mivggliglr
V7										
1579A	ivfavlsvin	rivrqqyspls	fqhtpnpgr	ldrpgriee	gqeqdrtrsi	rlvsgflala	wddrlrlclf	syhlrdlrfw	iaaktvllg	hsalqgllrg
CONSENSUS_A1	ivfavlsvin	rivrqqyspls	fqhtpnpgr	ldrpgriee	gqeqdrtrsi	rlvsgflala	wddrlrlclf	syhlrdlrfw	iaaktvllg	hsalqgllrg
CONSENSUS_A2	ivfavlsvin	rivrqqyspls	fqhtpnpgr	ldrpgriee	gqeqdrtrsi	rlvsgflala	wddrlrlclf	syhlrdlrfw	iaaktvllg	hsalqgllrg
Indie_C1	ivfavlsvin	rivrqqyspls	fqhtpnpgr	ldrpgriee	gqeqdrtrsi	rlvsgflala	wddrlrlclf	syhlrdlrfw	iaaktvllg	hsalqgllrg
V8										
1579A	weglylwnl	llywgllekk	sainlfdtia	ivagwtdri	ieigqgllgr	ilhiprrirq	glenfll-hg	1		
CONSENSUS_A1	weglylwnl	llywgllekk	sainlfdtia	ivagwtdri	ieigqgllgr	ilhiprrirq	glenfll-hg	1		
CONSENSUS_A2	weglylwnl	llywgllekk	sainlfdtia	ivagwtdri	ieigqgllgr	ilhiprrirq	glenfll-hg	1		
Indie_C1	weglylwnl	llywgllekk	sainlfdtia	ivagwtdri	ieigqgllgr	ilhiprrirq	glenfll-hg	1		

Fig. 5. Molecular characterization of HIV-1 p1579A-1 envelope gene. Amino acid alignment of p1579A-1, reference subtype A1 and A2 (subgroups of HIV-1 group M, subtype A) consensus sequences obtained from Los Alamos HIV-1 database and IndieC1 envelope gene. N—predicted N-linked glycosylation site. The GPGQ motif is surrounded by solid box. Positions 11 and 25 are represented by an asterisk (\*).

pIndieC1, as reported earlier (Mochizuki et al., 1999). An analysis of potential N-linked glycosylation sites did not show any major differences between p1579A-1 and reference subtype A1 or A2 obtained from the Los Alamos HIV database or pIndieC1. The number of predicted potential sites ranged from 27 to 31, with pIndieC1 having the most, 31 and p1579A-1 having 29. Variations in the locations and number of sites were seen predominantly in the variable regions V1, V2, V4, and V5. Only one potential site was predicted for p1579A-1 between positions 131 and 150 of the V1, region whereas at least 2 sites were predicted for other subtype A's and 3 were predicted for IndieC1 (Fig. 6B). Lack of potential N-linked glycosylation sites in this region is probably due to a 14-amino acid deletion in this region. Absence of glycans in this region could have an effect on CD4-gp120 interaction and viral entry (Chen et al., 2001; Kwong et al., 1998).

The long terminal repeat (LTR) is known to contain essential elements that control HIV-1 transcription and hence replication. As a first step to explore biological differences between subtypes from India, we compared the LTR of p1579A-1 with the LTR of pIndieC1 by performing a transcription factor binding site (TFBS) prediction analysis. Several differences were observed in the number of TFBS between subtypes A and C (Fig. 6). Most notably, both p1579A-1 and pIndieC1 were found to contain two NF-kappa B binding sites. This is in contrast to HIV-1 subtype C of African origin which has been shown to carry three functional NF-kappa B enhancers. This is not unusual, since Indian subtype C has been shown to be genetically distinct from other HIV-1 subtype C around the world (Shankarappa et al., 2001). Furthermore, pIndieC1 and

not p1579A-1 has the binding sites for GATA-binding factor 3 (GATA) and octamer-binding factor 1 (Oct1). GATA 3 is a member of a family of transcriptional activating proteins and is expressed in T lymphocytes. In vitro studies have shown enhanced HIV-1 LTR-directed transcription in the presence of this protein (Yang and Engel, 1993). Octamer-binding proteins (Oct) have been reported to both positively and negatively regulate the expression of a variety of genes (Zhang et al., 2004; Liu and Latchman, 1997). In vitro studies have shown Oct-1 to either repress LTR-directed transcription of HIV-1 in fibroblasts (Liu and Latchman, 1997) or to have no effect on HIV-1 transcription in primary CD4 T Cells (Zhang et al., 2004). Finally, p1579A-1 but not pIndieC1 has the binding site for GC Box elements which have been reported to be a component of Vpr-mediated LTR activation (Kuiken et al., 2001). It can be hypothesized that the presence or absence of these binding sites in the LTR of subtypes C and A may be responsible for differences in replication efficiency, resulting in the higher prevalence of subtype C over other subtypes of HIV-1 in India. Further studies are necessary to evaluate this hypothesis.

In conclusion, out of 19 HIV-1 Indian isolates, only one was found to be a subtype A. This finding is consistent with the low frequency (5–10%) of subtype A prevalence in India. From this HIV-1 subtype A isolate the molecular clone p1579A-1 was constructed. This clone was found to be replication competent in CD8 depleted PBMC and CCR5 tropic as determined by genotypic and phenotypic analysis. Sequence analysis of p1579A-1 indicate it to be phylogenetically linked to HIV-1 subtype A. p1579A-1 represents the first infectious molecular clone of HIV-1 subtype A which replicates in

## (A)

Transcription Factor Binding Site Predictions		1579A	Consensus A	Indie C1
<b>AHRARNT</b>	(Aryl hydrocarbon receptor/Arnt heterodimers)	1	0	0
<b>AML3</b>	(Runt-related transcription factor 2/CBFA19core binding factor, runt domain, alpha subunit 1))	1	1	0
<b>AP4</b>	(Activator protein 4)	1	1	0
<b>CMYB</b>	(c-Myb,important in hematopoiesis,cellular equivalent to avian myoblastosis virus oncogene v-myb)	1	1	1
<b>CP2</b>	(LBP-1c(leader-binding protein-1c), LSF(late SV40 factor), CP2, SEF(SAA3 enhanced factor)	1	1	1
<b>CREL</b>	(c-Rel)	0	0	1
<b>EVI1</b>	(Ecotropic viral integration site 1 encoded factor)	1	1	1
<b>GATA</b>	(GATA-binding factor 3)	0	0	1
<b>GC</b>	(GC box elements)	1	1	0
<b>MEIS1</b>	(Binding site for monomeric Meis1 homeodomain protein)	2	2	1
<b>NFAT</b>	(Nuclear Factor of activated T-cells)	0	0	1
<b>NFKAPPAB</b>	(NF-kappa B)	2	2	2
<b>OCT1</b>	(Octamer-binding factor 1)	0	1	1
<b>PAX6</b>	(PAX6 paired domain and homeodomain are required for binding to this site)	2	1	0
<b>SMAD3</b>	(Smad3 transcription factor involved in TGF-beta signaling)	0	0	1
<b>SP1</b>	(Stimulating protein 1 SP1, ubiquitous zinc finger transcription factor)	0	0	1
<b>VMAF</b>	(v-Maf)	1	0	1

## (B)

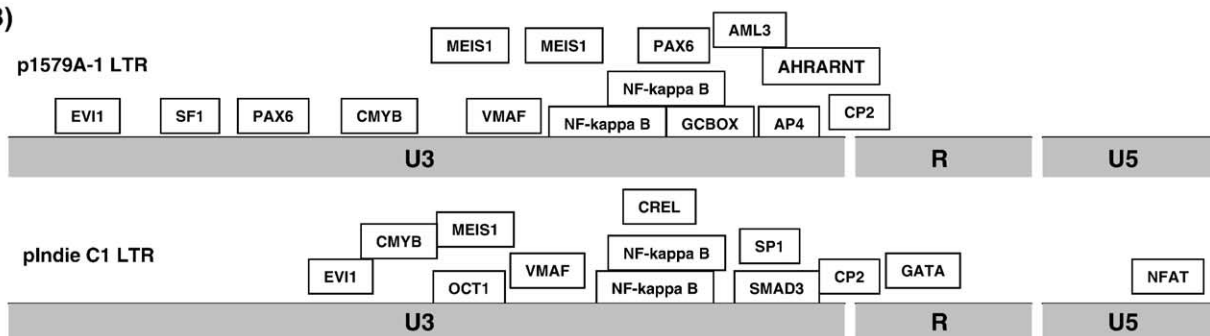


Fig. 6. Analysis of HIV-1 long terminal repeat (LTR). (A) Chart listing predicted transcription factor binding sites. Values in each column represent the number of sites predicted for each transcription factor binding site. (B) Schematic representation of predicted TFBS within the LTRs of 1579A (upper) and IndieC1 (lower).

primary CD4 lymphocytes. Availability of a molecular clone of subtype A HIV-1 will be useful for the study of replication dynamics and evolution of subtypes in India. Furthermore, its full genomic sequence will be important for characterizing other HIV-1 subtype A around the world.

## Materials and methods

### Virus cultures

PBMC from 19 HIV-1 positive Indian patients was cultured in the presence of phytohemagglutinin (PHA)-stimulated, CD8 depleted normal donor PBMC in RPMI 1640 media supplemented with IL-2 as described previously (Balachandran et al., 1988; Balachandran et al., 1991). Production of virus was monitored by HIV-1 p24 production in the culture supernatant. Virus was harvested every 5 days for 40 days. 10–20 million cultured cells were pelleted and stored at  $-80^{\circ}\text{C}$  for DNA isolation and subtyping.

### DNA isolation and HIV-1 env subtyping by HMA

High molecular weight chromosomal DNA was purified from cultured cell pellets using PUREGENE DNA purification kit. The DNA was then subjected to a heteroduplex mobility assay (HMA), as described previously (Delwart et al., 1993).

ED5/ED12 primers were used to PCR amplify the V1–V5 region of the envelope gene followed by a nested PCR amplification of the C2–V5 region using primers DR7/DR8. The cycling conditions for both primer pairs were as follows: 3 cycles of  $94^{\circ}\text{C}$  for 1 min,  $55^{\circ}\text{C}$  for 1 min,  $72^{\circ}\text{C}$  for 1 min, followed by 32 cycles of  $94^{\circ}\text{C}$  for 15 s,  $55^{\circ}\text{C}$  for 45 s,  $72^{\circ}\text{C}$  for 1 min. Final extension was at  $72^{\circ}\text{C}$  for 5 min. DR7/DR8 PCR products from each of the reference plasmids and the unknown HIV-1 were then combined, denatured at  $99^{\circ}\text{C}$  for 5 min, and then re-annealed on wet ice for 10 min. Heteroduplexes were resolved on a 5% non-denaturing polyacrylamide gel at 250 V for 3 h. The gel was then stained with ethidium bromide and visualized in a phosphorimager.

### Construction of the full-length infectious molecular clone

The entire 1579A genome was amplified in three fragments. Long PCR was performed using the high fidelity AccuPrime Pfx DNA polymerase (Invitrogen) and 1  $\mu\text{g}$  of chromosomal DNA from HIV-1 1579A infected PBMC. Three overlapping subgenomic fragments of 2.5 kb, 5'/LTR-gag, 6.2 kb, gag-env, and 2.7 kb, env-LTR3', were generated using PCR primers shown in Table 3. The cycling conditions were as follows: 1 cycle of  $95^{\circ}\text{C}$  for 2 min. Followed by 10 cycles of  $94^{\circ}\text{C}$  for 10 s,  $55^{\circ}\text{C}$  for 30 s and  $66^{\circ}\text{C}$  for 2 min, followed by 25 cycles of  $94^{\circ}\text{C}$  for 15 s,  $55^{\circ}\text{C}$  for 30 s and  $66^{\circ}\text{C}$  for 2 min. Final

Table 3  
Primers used for amplification of HIV-1 subtype A genome

Fragment length (kb)	Position	Primer	Sequence (5' → 3')
2.5	5'LTR-pol	U3XhoI	ATT ACT CGA GTG GAT GGG TTA ATT
		HIV-1PolI	ACT GGT ACA GTC TCA ATA GGA CTA ATT G
6.2	gag-env	SK145	AGT GGG GGG ACA TCA AGC AGC CAT GCA AAT
		DR8	GGG ACA ATT GGA GAA GTG
		DR7	CAA CTG CTG TTA AAT GGC AGT CTA GC
2.7	env-3'LTR	New LTR	ATA AGC GGC CGC CCA
		3' NotI	CTG CTA GAG ATT T

extension was at 66 °C for 7 min. Extension times were increased by 5 s every additional cycle. Annealing temperatures for 2.5 kb, 6.2 kb and 2.7 kb were 50 °C, 55 °C and 57 °C, respectively. PCR products were gel purified using Mo Bio Laboratories DNA purification kit and then cloned into pCR-Blunt II TOPO vectors (Invitrogen) resulting in plasmids A, B, and C, respectively. The plasmid A carrying the 2.5-kb HIV-1 DNA fragment and the plasmid B carrying the 6.2-kb HIV-1 DNA fragment were digested with unique restriction enzymes. The 4.4-kb fragment from the plasmid A was ligated to the 6.7-kb fragment from the plasmid B to generate a plasmid D that contained a 7.6-kb fragment encompassing the 5'LTR-env region of the 1579A HIV-1 genome. The plasmid C carrying the 2.7-kb HIV-1 DNA fragment and plasmid D were then digested with unique restriction enzymes. The 9.6-kb fragment from the plasmid D was ligated to the 3.9-kb fragment from the plasmid C to generate a plasmid that contained the complete full-length HIV-1 proviral genome of 1579A and was designated p1579A.

#### *Transfection, replication kinetics, coreceptor usage, and MT2 assay*

HEK293T cells (ATCC) were maintained in DMEM supplemented with 15% FBS, 1% L-glutamine, 1% penicillin–streptomycin, 1 µg/ml puromycin and 300 µg/ml G418. 293T cells that were grown to 90–95% confluency were transfected with 14 µg of p1579A-1 DNA or 35 µg of control plasmid pIndieC1 with Lipofectamine 2000 reagent (Invitrogen). Transfected HEK293T cells were cultured in 15% DMEM media for 48 h. After 48 h, supernatant from transfection cultures was passed through a 0.45 µm filter and added to 5 million PHA stimulated CD8 depleted PBMC treated with 5 µg/ml polybrene for 1 h in RPMI media supplemented with IL-2. Cultures were maintained at 37 °C. Cultures were tested for the presence of p24 antigen every 2–3 days and fed with 2.5 million PHA stimulated and polybrene treated CD8 depleted PBMC every 7 days.

To determine the coreceptor usage of p1579A-1 virus, U87.CD4 cells (NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH) were infected as described (Bjorndal et al., 1997). Briefly, U87.CD4 cells

expressing either the CXCR4 or CCR5 chemokine receptors were seeded in duplicate in a 24-well microtiter plate at a concentration of 250,000 cells/well in a final volume of 500 µl 15%D-MEM. The cells were incubated at 37 °C until they became 90% confluent. The media were removed from confluent cells, and the cell monolayer was washed with 500 µl PBS. Approximately 0.5 ml (30 pg p24 equivalent) of p1579A-1 or the control HIV-1 BAL or IIIB was added to duplicate wells of U87.CD4.CCR5 and CXCR4 cells. One set of wells for each coreceptor was left uninfected as a negative control. An additional 0.5 ml of DMEM containing 15% FCS was added to each plate, and the plates were incubated at 37 °C for 24 h. After 24 h, the supernatant was removed. The cells were rinsed with PBS and incubated in 1 ml fresh DMEM supplemented with 15% FCS. HIV-1 p24 production in the supernatant was monitored every 2 days.

To determine the replication kinetics of p1579A-1 cloned virus 10 million PHA-stimulated CD8 depleted PBMC treated with 5 µg/ml polybrene for 1 h were infected with approximately 30 ng equivalent HIV-1 p24 virus of either cloned virus or the parental isolate. The infections were carried out for 2 h, with shaking every 30 min at 37 °C. After the 2 h, the virus supernatant was removed, and the cells were washed and cultured at 37 °C in RPMI 1640 medium containing 20% FCS and 5% IL-2. Virus growth was monitored every 2 days by p24 antigen production in the culture supernatant.

MT2 assay to measure syncytia inducing activity of HIV-1 was performed as described previously (O'Marino et al., 1992).

#### *Sequencing, phylogenetic analysis and molecular characterization*

The entire viral genome was sequenced using an automated sequencer. Sequences were manually joined and contiguous sequence fragments were assembled using Vector NTI suite (Informax, Oxford, UK). DNA and protein alignments were constructed using CLUSTAL X alignment program of Vector NTI. Consensus sequences and subtype reference sequences used for alignments were obtained from the Los Alamos HIV-1 database.

Phylogenetic analyses of the entire genome sequence were constructed by the neighbor-joining method of Jukes Cantor corrected distances (protein, total mean character corrected distances) with the optimality criterion set to distance as measured in PAUP. Statistical significance of branchings and various clustering were assessed by bootstrap re-sampling of 1000 pseudoreplicates on the complete data set. The trees were edited for publication using Treeview68K version 1.5. Phylogenetic analyses of LTR sequences were conducted through the construction of maximum likelihood trees as measured in PAUP. Statistical significance of branchings and various clustering were assessed by bootstrap re-sampling of 1000 pseudoreplicates on the complete data set. The trees were edited for publication using Treeview68K version 1.5.



N-glycosylation site analysis was performed at the Los Alamos HIV-1 database. Transcription factor binding site predictions were performed using MatInspector.

Accession Number: DQ083238

## Acknowledgments

We thank Ming Ding, Kathy Kulka, and Mary White for their technical assistance. This work was supported by AIDS-FIRCA grant R03 TH00971 (PG) and AI51661 (SH).

## References

- Bachmann, M.H., Delwart, E.L., Shpaer, E.G., Lingenfelter, P., Singal, R., Mullins, J.I., 1994. Rapid genetic characterization of HIV type 1 strains from four World Health Organization-sponsored vaccine evaluation sites using a heteroduplex mobility assay. WHO Network for HIV Isolation and Characterization. *AIDS Res. Hum. Retroviruses* 10 (11), 1345–1353.
- Balachandran, R., Thampatty, P., Rinaldo, C., Gupta, P., 1988. Use of cryopreserved normal peripheral blood lymphocytes for isolation of human immunodeficiency virus from seropositive men. *J. Clin. Microbiol.* 26 (3), 595–597.
- Balachandran, R., Thampatty, P., Enrico, A., Rinaldo, C., Gupta, P., 1991. Human immunodeficiency virus isolates from asymptomatic homosexual men and from AIDS patients have distinct biologic and genetic properties. *Virology* 180 (1), 229–238.
- Bjorndal, A., Deng, H., Jansson, M., Fiore, J.R., Colognesi, C., Karlsson, A., Albert, J., Scarlatti, G., Littman, D.R., Fenyo, E.M., 1997. Coreceptor usage of primary human immunodeficiency virus type 1 isolates varies according to biological phenotype. *J. Virol.* 71 (10), 7478–7487.
- Cassol, S., Weniger, B.G., Babu, P.G., Salminen, M.O., Zheng, X., Htoon, M.T., Delaney, A., O'Shaughnessy, M., Ou, C.Y., 1996. Detection of HIV type 1 env subtypes A, B, C, and E in Asia using dried blood spots: a new surveillance tool for molecular epidemiology. *AIDS Res. Hum. Retroviruses* 12 (15), 1435–1441.
- Cecilia, D., Kulkarni, S.S., Tripathy, S.P., Gangakhedkar, R.R., Paranjape, R.S., Gadkari, D.A., 2000. Absence of coreceptor switch with disease progression in human immunodeficiency virus infections in India. *Virology* 271, 253–258.
- Chen, M., Shi, C., Kalia, V., Tencza, S.B., Montelaro, R.C., Gupta, P., 2001. HIV gp120 V(1)/V(2) and C(2)-V(3) domains glycoprotein compatibility is required for viral replication. *Virus Res.* 79 (1–2), 91–101.
- Delwart, E.L., Shpaer, E.G., Louwagie, J., McCutchan, F.E., Grez, M., Rubsamen-Waigmann, H., Mullins, J.I., 1993. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. *Science* 262 (5137), 1257–1261.
- Delwart, E.L., Herring, B., Rodrigo, A.G., Mullins, J.I., 1995. Genetic subtyping of human immunodeficiency virus using a heteroduplex mobility assay. *PCR Methods Appl.* 4 (5), S202–S216.
- Delwart, E.L., Mullins, J.I., Gupta, P., Learn Jr., G.H., Holodniy, M., Katzenstein, D., Walker, B.D., Singh, M.K., 1998. Human immunodeficiency virus type 1 populations in blood and semen. *J. Virol.* 72, 617–623.
- Jameel, S., Zafrullah, M., Ahmad, M., Kapoor, G.S., Sehgal, S., 1995. A genetic analysis of HIV-1 from Punjab, India reveals the presence of multiple variants. *AIDS* 9 (7), 685–690.
- John, T.J., Babu, P.G., Jayakumari, H., Simoes, E.A., 1987. Prevalence of HIV infection in risk groups in Tamilnadu, India. *Lancet* 1 (8525), 160–161.
- Kalia, V., Sarkar, S., Gupta, P., Montelaro, R.C., 2003. Rational site-directed mutations of the LLP-1 and LLP-2 lentivirus lytic peptide domains in the intracytoplasmic tail of human immunodeficiency virus type 1 gp41 indicate common functions in cell–cell fusion but distinct roles in virion envelope incorporation. *J. Virol.* 77 (6), 3634–3646.
- Kuiken, C.L., Foley, B., Hahn, B., Korber, B., Marx, P.A., McCutchan, F., Mellors, J.W., Mullins, J.I., Sodroski, J., Wolinsky, S. (Eds.), 2001. HIV Sequence Compendium. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, LA-UR 01-3860.
- Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., Hendrickson, W.A., 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393 (6686), 648–659.
- Liu, Y.Z., Latchman, D.S., 1997. The octamer-binding proteins Oct-1 and Oct-2 repress the HIV long terminal repeat promoter and its transactivation by Tat. *Biochem. J.* 322 (Pt. 1), 155–158.
- Ljungberg, K., Hassan, M.S., Islam, M.N., Siddiqui, M.A., Aziz, M.M., Wahren, B., Islam, K.B., Leitner, T., 2002. Subtypes A, C, G, and recombinant HIV type 1 are circulating in Bangladesh. *AIDS Res. Hum. Retroviruses* 18 (9), 667–670.
- Maitra, A., Singh, B., Banu, S., Deshpande, A., Robbins, K., Kalish, M.L., Broor, S., Seth, P., 1999. Subtypes of HIV type 1 circulating in India: partial envelope sequences. *AIDS Res. Hum. Retroviruses* 15, 941–944.
- Mandal, D., Jana, S., Panda, S., Bhattacharya, S., Ghosh, T.C., Bhattacharya, S.K., Chakrabarti, S., 2000. Distribution of HIV-1 subtypes in female sex workers of Calcutta, India. *Indian J. Med. Res.* 112, 165–172.
- Mandal, D., Jana, S., Bhattacharya, S.K., Chakrabarti, S., 2002. HIV type 1 subtypes circulating in eastern and northeastern regions of India. *AIDS Res. Hum. Retroviruses* 18, 1219–1227.
- Mochizuki, N., Otsuka, N., Matsuo, K., Shiino, T., Kojima, A., Kurata, T., Sakai, K., Yamamoto, N., Isomura, S., Dhole, T.N., Takebe, Y., Matsuda, M., Tatsumi, M., 1999. An infectious DNA clone of HIV type 1 subtype C. *AIDS Res. Hum. Retroviruses* 15 (14), 1321–1324.
- O'Marro, S.D., Armstrong, J.A., Asuncion, C., Gueverra, L., Ho, M., 1992. The effect of combination of ampicillin and zidovudine or dideoxyinosine against human immunodeficiency viruses in vitro. *Antiviral Res.* 17, 169–177.
- Sahni, A.K., Prasad, V.V., Seth, P., 2002. Genomic diversity of human immunodeficiency virus type-1 in India. *Int. J. STD AIDS* 13 (2), 115–118.
- Shankarappa, R., Chatterjee, R., Learn, G.H., Neogi, D., Ding, M., Roy, P., Ghosh, A., Kingsley, L., Harrison, L., Mullins, J.I., Gupta, P., 2001. Human immunodeficiency virus type 1 env sequences from Calcutta in eastern India: identification of features that distinguish subtype C sequences in India from other subtype C sequences. *J. Virol.* 75 (21), 10479–10487.
- Simoes, E.A., Babu, P.G., John, T.J., Nirmala, S., Solomon, S., Lakshminarayana, C.S., Quinn, T.C., 1987. Evidence for HTLV-III infection in prostitutes in Tamil Nadu (India). *Indian J. Med. Res.* 85, 335–338.
- Tripathy, S., Renjifo, B., Wang, W.K., McLane, M.F., Bollinger, R., Rodrigues, J., Osterman, J., Essex, M., 1996. Envelope glycoprotein 120 sequences of primary HIV type 1 isolates from Pune and New Delhi, India. *AIDS Res. Hum. Retroviruses* 12 (12), 1199–1202.
- Tsuchie, H., Saraswathy, T.S., Sinniah, M., Vijayamalar, B., Maniar, J.K., Monzon, O.T., Santana, R.T., Paladin, F.J., Wasi, C., Thongcharoen, P., et al., 1995. HIV-1 variants in South and South-East Asia. *Int. J. STD AIDS* 6 (2), 117–120.
- UNAIDS, W.H.O., 2004. Report on the Global AIDS Epidemic. UNAIDS Report. UNAIDS, Geneva, Switzerland.
- Yang, Z., Engel, J.D., 1993. Human T cell transcription factor GATA-3 stimulates HIV-1 expression. *Nucleic Acids Res.* 21 (12), 2831–2836.
- Zhang, M., Genin, A., Cron, R.Q., 2004. Overexpression of octamer transcription factors 1 or 2 alone has no effect on HIV-1 transcription in primary human CD4 T cells. *Virology* 321 (2), 323–331.